Ungraded Problems Assigned in Week 7 (Chapter 5)

Math 394

#3

The scheme is exactly the same as in problem 1: we need to make sure that the function is nonnegative, and integrable. Now

$$2x - x^3 \geq 0$$

if $x \geq 0$ and $x^2 \leq 2$, i.e., in $[0, \sqrt{2}]$. Since $\frac{5}{2} > \sqrt{2}$ this is not a candidate for probability density.

Similarly,

$$2x - x^2 > 0$$

if $0 \leq x$, and $2 \geq x$. But $\frac{5}{2} = 2.5 > 2$ and this function is also not a suitable candidate.

#6

Here, we use the formula

$$EX = \int_{-\infty}^{\infty} x f_X(x) dx$$

which is meaningful only if xf_X is an integrable function (for instance, a density of the form $\frac{c}{x^2}$ for $x \geq 1$, with suitable c, would be indeed a density, but it would have no expected value. More to the point, the Cauchy distribution (with density of the form $\frac{c}{a^2+x^2}$), which appears in several applications, also has no expected value).

a First, note that f is indeed a density function (it is a member of the family of so-called "Gamma" distributions) :

$$\int_0^{\infty} \frac{1}{4} x e^{-\frac{x}{2}} dx = \frac{1}{4}\left(\left[-2xe^{-\frac{x}{2}}\right]_0^{\infty} + 4\right) = 1$$

To compute EX, we calculate

$$\int_{-\infty}^{\infty} x f(x) dx = \frac{1}{4}\int_0^{\infty} x^2 e^{-\frac{x}{2}} dx$$

by integrating by parts twice. To streamline the calculations, it doesn't hurt to remember (from the exponential distribution - cf. p.209 in the book) that

$$\int_0^\infty ke^{-kx}dx = 1, \quad \int_0^\infty kxe^{-kx}dx = \frac{1}{k}, \quad \int_0^\infty kx^2e^{-kx}dx = \frac{2}{k^2}$$

It follows that $EX = 4$

b We already saw in problem 1 that $c = \frac{3}{4}$. As for the expected value,

$$EX = \frac{3}{4}\int_{-1}^{1} x\left(1-x^2\right)dx = \frac{3}{4}\left[\frac{x^2}{2} - \frac{x^2}{4}\right]_{-1}^{1} = 0$$

Of course, this could have been seen immediately, since we are integrating an odd function ($f(x) = -f(-x)$) over an interval symmetric around 0.

c First of all, this function is a density, since

$$\int_5^\infty \frac{5}{x^2} = \left[-\frac{5}{x}\right]_5^\infty = 1$$

but it has no expected value: $x \cdot \frac{5}{x^2} = \frac{5}{x}$ is notoriously not integrable at $+\infty$ ($\int_1^X \frac{5}{x}dx = \log X \to \infty$ as $X \to \infty$).

#7

A we noted in problem 1, a density function has to satisfy $f \geq 0$, and $\int_{-\infty}^{\infty} f dx = 1$. The last formula provides one relation to determine the missing constants. To determine both we need another relation, and that's provided by the value of $\int_{-\infty}^{\infty} xf dx$. We have two equations in two unknowns a and b:

$$\int_0^1 \left(a + bx^2\right)dx = a + \frac{b}{3} = 1$$

$$\int_0^1 x\left(a + bx^2\right)dx = \frac{a}{2} + \frac{b}{4} = \frac{3}{5}$$

Solving this simple system gives $a = \frac{3}{5}, b = \frac{6}{5}$

#9

Let's go to example 4b on p.130 of the book. Recapping, we consider a random variable X representing the number of units of a product *sold* by a department store that has to stock them in advance[1]. Suppose a stock of s has been ordered. Then a profit of b per unit sold, and a loss of l per unit left unsold

[1]The book, mysteriously, talks about X as the amount of units "ordered", rather than sold. Maybe they were thinking of a mail-order only store, but it sure makes for a confusing discussion. In any case, equation (1) (which is also the equation at the bottom of p. 134) makes clear what this is all about.

is recorded. Incidentally, we are not considering "virtual losses", due to missed sales because of insufficient stocked items that cause some demand not to be satisfied. The equation is then (cf. the first equation on page 130), calling the profit, a function of s, $P(s)$,

$$P(s) = b(s \wedge X) - (s - X)^{+} l \tag{1}$$

(we introduced a couple of common notations: $a \wedge b = \min(a, b)$, and

$$a^{+} = \begin{cases} a & a > 0 \\ 0 & a \le 0 \end{cases} = \max(a, 0)$$

to shorten the writing).

Now, the expected profit is

$$EP(s) = bE[s \wedge X] - lE(s - X)^{+} \tag{2}$$

and some care is needed, since neither function of X is linear!

As you may guess by looking at the discrete variable solution on p. 135, the trick is to break up the expectation in two parts, depending on whether $s < X$ or not. Equivalently, we could introduce two RVs

$$Y = s \wedge X; Z = (s - X)^{+}$$

and write down their densities (actually, they are "mixed" random variable: they take a specific value with a positive probability, while the remaining values are continuous:

$$P[Y = s] = P[X \ge s]; P[Z = 0] = P[X \ge s]$$

and a density elsewhere) while, conditioning on $X \le s$, we have a (conditional) density, the density of X, suitably normalized.

There are tricks to represent even the distribution of discrete RVs in terms of "densities" (recall that some books refer to the pmf of a discrete RV as a "discrete density"). This is done in a handwaving manner by introducing the so called "$\delta-$function" (which is not a function at all) $\delta(x - a)$, where δ is characterized by the property that, for any continuous function,

$$\int_{-\infty}^{\infty} f(x)\, \delta(x - a)\, dx = f(a)$$

which is sometimes "explained" in physics and engineering books by stating that $\delta$is "equal to zero everywhere, except at a, where it is infinite" (there are rigorous mathematical ways to define such an object - as a Schwartz "distribution" - no relation with probability distributions, of course - or as a measure, or as a linear functional on an appropriate function space - all the preceding terms are, in a sense, equivalent or closely connected). Hence, a pmf $p_1, p_2, \ldots, p_n$, for values $x_1, x_2, \ldots, x_n$ can be written (formally) as

$$f(x) = \sum_{k=1}^{n} p_k \delta(x - x_k) \tag{3}$$

This is only meant to show that the following solution for the problem, assuming X is a continuous RV, applies to the discrete (or even mixed) case as is, using the formal trick of equation (3).

We compute the expected values in (2):

$$E[s \wedge X] = sP[X > s] + \int_0^s x f(x)\, dx$$

(here f is the density of X)

$$E\left[(s - X)^+\right] = \int_0^s (s - x) f(x)\, dx$$

so that the expected profit is

$$bsP[X > s] + b\int_0^s x f(x)\, dx - lsP[X \leq s] + l\int_0^s x f(x)\, dx$$

Let's derive this expression with respect to s (noting that $P[X > s] = 1 - F(s)$, and $[X \leq s] = F(s)$, where F is the cdf of X, and that $F' = f$):

$$b(1 - F(s)) - bsf(s) + bsf(s) - lF(s) - lsf(s) + lsf(s) =$$

$$= b - (b + l) F(s)$$

which is zero for any s^* such that

$$F(s^*) = \frac{b}{b + l} \tag{4}$$

Note that there could be more than one s^*, if the function F happened not to be strictly monotone. However, for instance in the discrete case, only the left endpoint of the interval of s satisfying (4) would have physical meaning (if you are counting single units, you cannot sell "2.45 single units"). This has been expressed as an "inequality" condition in the solution to the discrete problem

#10

a We have a density function

$$f_p(t) = \begin{cases} 1 & 7 \leq t \leq 8 \\ 0 & \text{otherwise} \end{cases}$$

for the arrival time of the passenger. The passenger will take the train to B if he/she arrives between 7 and $7:05$, $7:15$ and $7:20$, and so on, up to

7 : 45 to 7 : 50, for a total time "window" of 20 minutes. If he/she arrives in the remaining time intervals, the "winning" train goes to A, and that is over a total "window" of 40 minutes. Hence, the probability of going to A is $\frac{40}{60} = \frac{2}{3}$. Incidentally, the probability of the passenger catching the train that leaves at 7:00 is zero, since $\int_0^0 f_p(t)dt = 0$.

b At first, we have 5 minutes of opportunity to get the next train, which now is the one going to A. After this initial period, the passenger has again 10-minute windows to catch a train to A, alternating with 5-minute windows to catch a train to B, and, if late, will catch the 8:00 train to A. Hence the percentage for the two destinations is the same as before: the probability of going to A is $\frac{2}{3}$.

#19

This type of problem works by recasting the question in terms of a *standard* normal distribution, and then going to the tables. Thus,

$$P[X > c] = P\left[\frac{X - EX}{\sqrt{Var[X]}} > \frac{c - EX}{\sqrt{Var[X]}}\right] = P\left[\frac{X - EX}{\sqrt{Var[X]}} > \frac{c - 12}{\sqrt{4}}\right] =$$

$$P\left[Z > \frac{c}{2} - 6\right]$$

where Z is a standard normal. Reading tables in reverse, we find out that

$$P[Z > z_{.1}] = .10$$

where $z_{.1} = 1.2816$. Hence,

$$\frac{c}{2} = 7.2816$$
$$c = 14.5632$$

#23

The number 6 appears with probability $\frac{1}{6}$, thus the number of appearances over n throws is a binomial with parameters $n, \frac{1}{6}$. If X is this random number,

$$\frac{X - \frac{n}{6}}{\sqrt{\frac{n}{6} \cdot \frac{5}{6}}}$$

is approximately a standard normal. For $n = 1000$, $150 \le X \le 200$ if

$$\frac{150 - \frac{1000}{6}}{\sqrt{\frac{1000}{6} \cdot \frac{5}{6}}} \le \frac{X - \frac{1000}{6}}{\sqrt{\frac{1000}{6} \cdot \frac{5}{6}}} \le \frac{200 - \frac{1000}{6}}{\sqrt{\frac{1000}{6} \cdot \frac{5}{6}}}$$

i.e.

$$\frac{900 - 1000}{\sqrt{5000}} \le \frac{X - \frac{1000}{6}}{\sqrt{\frac{1000}{6} \cdot \frac{5}{6}}} \le \frac{1200 - 1000}{\sqrt{5000}}$$

$$-\frac{100}{50\sqrt{2}} \le Z \le \frac{200}{50\sqrt{2}}$$

where Z is (approximately) a standard normal. Since

$$P\left[-\sqrt{2} \leq Z \leq 2\sqrt{2}\right] = .91091$$

this is the requested probability[2].

On the other hand, if 6 appeared exactly 200 times, the remaining 800 throws were divided between the remaining 5 possible outcomes. Out of 800 throws, the number of 5s (conditional probability $\frac{1}{5}$) Y satisfies

$$P[Y < 150] = P\left[\frac{Y - 800 \cdot \frac{1}{5}}{\sqrt{800 \cdot \frac{1}{5} \cdot \frac{4}{5}}} < \cdot \frac{750 - 800}{2 \cdot \sqrt{800}}\right] =$$

$$= P\left[Z < -\frac{5}{4\sqrt{2}}\right] = .18838$$

Remark The solution manual suggests to calculate $P[149.5 < X < 200.5]$, instead of what we did above, i.e., $P[150 \leq X \leq 200]$. This is a homage to a practice (sometimes called "smoothing") that purports to take into account the fact that we are approximating a discrete distribution with a continuous one, by, somehow, interpolating by halves between the "really possible" values. This practice is not completely justified. In fact, if the number of trials is too small to invoke the Central Limit Theorem straight, it is unlikely that this correction will go in the right direction anyway. If the number of trials is adequate (and for a binomial distribution, with p not too close to 0 or to 1, such a number is really low), there is simply no point in doing so, since $\frac{.5}{\sqrt{n}}$ will in any case be of little relevance. Also, since we are doing an approximation anyway, it doesn't make too much sense to push the number of significant digits too far, which additionally makes the correction more a matter of whim that a sound mathematical need.

#31

You might recall the following fact: for any RV X, the number m that minimizes the function $\phi(m) = E\left[(X-m)^2\right]$ is $m = EX$. This is easily proved, by just expanding the expression, and observing that a parabola concave up is lowest at its vertex. This can be expressed by saying that "the number that minimizes the quadratic expected cost is the expected value", and provides a connection between the variance (and the idea of "quadratic cost"), and the expected value.

Another way of "measuring cost" is the one put forward here: minimize the expected ***absolute value*** of the error, instead of its square. Let's try to answer the question for a general continuous distribution (the argument for a discrete distribution is even more straightforward): let X have density f_X, and cdf F_X. We look for a number a that minimizes

$$\psi(a) = E[|X-a|] = \int_{-\infty}^{\infty} |x-a| f_X(x)\,dx$$

Due to the absolute value, there is no way we can solve this by deriving with respect to a - which illustrates why it is so much nicer to work with quadratic costs! By definition,

$$|x-a| = \begin{cases} x-a & x \geq a \\ a-x & x \leq a \end{cases}$$

so that

$$\psi(a) = \int_{-\infty}^{a} (a-x) f_X(x)\,dx + \int_{a}^{\infty} (x-a) f_X(x)\,dx =$$

$$= a\left(\int_{-\infty}^{a} f_X dx - \int_{a}^{\infty} f_X dx\right) - \left(\int_{-\infty}^{a} x f_X dx - \int_{a}^{\infty} x f_X dx\right)$$

This expression we can derive with respect to a in an easy way:

$$\psi'(a) = \int_{-\infty}^{a} f_X dx - \int_{a}^{\infty} f_X dx + 2a f_X(a) - 2a f_X(a) =$$

$$F_X(a) - (1 - F_X(a))$$

and this is zero when

$$2F_X(a) = 1$$

or

$$F_X(a) = P[X \leq a] = \frac{1}{2}$$

This value is called the "median" of the distribution (it represent the value for which X has probability 0.5 of being either lower or higher), and it is often used in describing empirical data (e.g., newspapers will report the "median price of housing", the "median income", etc. in a given area).

For our specific questions, we thus have that

a a is the median of the uniform distribution on $[0, A]$, and it is obvious that $a = \frac{A}{2}$

b a is the median of the exponential distribution, and this is the value such that

$$P[X \leq a] = P[X > a] = 0.5$$

i.e.

$$\begin{aligned} e^{-\lambda a} &= 0.5 \\ -\lambda a &= \log\left(\frac{1}{2}\right) \\ a = \frac{\log(2)}{\lambda} &= \tau \cdot \log(2) \end{aligned}$$

Note the connection between a and the expected value $\tau = \frac{1}{\lambda}$. It is the median that is often called "the half life" of a radioactive material, whose decay law is assumed to follow an exponential distribution. It is almost the same as the expected value, up to a factor of $\log(2) = .69315$.

Of course, we obtain the same result if we work directly with the given distributions, without referring to the general result.

a We look for a that minimizes

$$\psi(a) = \int_0^A |x-a| \frac{1}{A} dx = \frac{1}{A}\left(\int_0^a (a-x)\,dx + \int_a^A (x-a)\,dx\right) =$$

$$= \frac{1}{A}\left(a^2 - \frac{a^2}{2} + \left(\frac{A^2-a^2}{2}\right) - a\,(A-a)\right) = \frac{1}{A}\left(\frac{A^2}{2} - aA + a^2\right)$$

which, as a function of a, is a parabola, concave up, with a minimum at its vertex, i.e. at $a = -\frac{-A}{2} = \frac{A}{2}$

b We now look at minimizing

$$\int_0^\infty \lambda\,|x-a|\,e^{-\lambda x} dx = \lambda\left(\int_0^a (a-x)\,e^{-\lambda x} dx + \int_a^\infty (x-a)\,e^{-\lambda x} dx\right) =$$

$$= \lambda\left(a\left(\int_0^a e^{-\lambda x} dx - \int_a^\infty e^{-\lambda x} dx\right) - \left(\int_0^a x e^{-\lambda x} dx + \int_a^\infty x e^{-\lambda x} dx\right)\right)$$

Multiplying throughout the first factor λ, we get

$$a - \frac{1}{\lambda}\left(1 - 2e^{-\lambda a}\right)$$

(we have used the cdf of an exponential, $1 - e^{-\lambda x}$, as well as integrated by parts the terms of the form

$$\int_c^d x e^{-\lambda x} dx = \left[-\frac{1}{\lambda} x e^{-\lambda x}\right]_c^d + \frac{1}{\lambda}\int_c^d e^{-\lambda x} dx =$$

$$= \frac{1}{\lambda}\left(ce^{-\lambda c} - de^{-\lambda d}\right) + \frac{1}{\lambda^2}\left(e^{-\lambda c} - e^{-\lambda d}\right)$$

substituting appropriately).
This function of a will be minimized at $1 + 2e^{-\lambda a} = 1$, or $a = \frac{\log 2}{\lambda}$

#36

A distribution with hazard function $h(t)$ is such that

$$P[X > t] = e^{-h(t)}$$

hence in our case,

a $P[X > 2] = e^{-\int_0^2 t^3 dt} = e^{-\frac{2^4}{4}} = e^{-4} = .018316$

b $P[.4 < X < 1.4] = e^{-\int_0^{0.4} t^3 dt} - e^{-\int_0^{1.4} t^3 dt} = e^{-\frac{.4^4}{4}} - e^{-\frac{1.4^4}{4}} = .61088$

c $P[X > 2 \,|X > 1] = \frac{P[X>2]}{P[X>1]} = e^{-\int_1^2 t^3 dt} = e^{-4+\frac{1}{4}} = .023518$

#37

As usual, let's be careful with absolute values.

a $P\left[|X| > \frac{1}{2}\right] = P\left[X > \frac{1}{2}\right] + P\left[X < -\frac{1}{2}\right] = 2 \cdot \frac{1}{4} = \frac{1}{2}$

b To find the density, we can proceed in several ways. Let's, for instance, compute the cdf (obviously, $F_X = 0$ for $x < 0$, so we restrict to $x > 0$):

$$P[|X| \le x] = P[-x \le X \le x] = \int_{-x}^{x} \frac{1}{2} dt = x$$

Hence, $|X|$ is uniform on $[0, 1]$.

Theoretical Exercises

#1

This is, really, a Calculus exercise. We are enforcing

$$\int_0^\infty ax^2 e^{-bx^2} = 1$$

The following method relies on the identity $\Gamma\left(\frac{1}{2}\right) = \sqrt{\pi}$, which you may be familiar with, or not (it is proved on page 199 in the book), and the property of the Gamma function (see page 215): $\Gamma(x) = (x-1)\Gamma(x-1)$

Taking this identity for granted, we can use substitution, namely $x^2 = u$, to transform the integral into

$$\frac{a}{2}\int_0^\infty \frac{\sqrt{b}u^{\frac{1}{2}}}{\sqrt{b}} e^{-bu} du = \frac{a\Gamma\left(\frac{3}{2}\right)}{b\sqrt{b}}\int_0^\infty \frac{b(bu)^{\frac{3}{2}-1}}{\Gamma\left(\frac{3}{2}\right)} e^{-bu}$$

From the definition of the Gamma distribution (on page 237), we know that the integral is equal to 1 (read b as λ in the formula). Hence, we have

$$a\frac{\Gamma\left(\frac{3}{2}\right)}{b^{\frac{3}{2}}} = 1$$

which, using the fact that $\Gamma\left(\frac{3}{2}\right) = \frac{1}{2}\Gamma\left(\frac{1}{2}\right) = \frac{\sqrt{\pi}}{2}$ results in

$$a = \frac{2b^{\frac{3}{2}}}{\sqrt{\pi}} = 2b\sqrt{\frac{b}{\pi}}$$

#2

Following the hint, we have (for absolutely continuous random variables)

$$E[Y] = \int_{-\infty}^\infty y f_Y(y) dy = \int_{-\infty}^0 y f_Y(y) dy + \int_0^\infty y f_Y(y) dy$$

Integrating by parts both integrals results in

$$[yF_Y(y)]_{-\infty}^\infty - \int_{-\infty}^\infty F_Y(y) dy$$

where F_Y is the cumulative distribution function of Y. To prove the assertions we only need to observe that the integral is precisely the right hand side of the equation (when a density function exists, the definition of cumulative distribution function in terms of < or $\leq$ makes no difference), and that, in order for the expected value to exist, the first term in the expression has to be zero (that is equivalent to the requirement that the integral $\int_{-\infty}^\infty y f_Y(y) dy$ exist!)

#3

The hint shows a pretty straightforward path. A less rigorous argument (that can, however, be extended to more general situations, once the necessary technicalities are settled) could go as follows.

Let X be a continuous random variable. We approximate it by X_n, where the range of X has been "cut up" in small segments $[a_k, a_{k+1})$, of length $\frac{1}{n}$, and $X_n = a_k$ for all outcomes such that $a_k \leq X < a_{k+1}$. We now have a discrete random variable, and we have, from Chapter 4

$$E[g(X_n)] = \sum_k g(a_k)P[X_n = a_k]$$

By the fact that $P[X_n = a_k] = \int_{a_k}^{a_{k+1}} f(x)dx$, and that the right hand side is now, essentially, a Riemann sum for our "target" integral, as $n \to \infty$, we have our proof.

Remark: the argument would work for *any* random variable X, if we had a rigorous method for passing to the limit in $\sum_k g(a_k)P[X \in [a_k, a_{k+1})]$ even when the distribution does not admit a continuous density. In fact, we would need to extend our definition of integral, in order for this to work, and that is precisely what the Lebesgue integration theory does for functions if a real variables, and general measure theory does for more general situations (recall that our X has an abstract space, the sample space, as its domain).

#5

Let's follow the hint:

$$E[X^n] = \int_0^\infty P[X^n > x^n] n x^{n-1} dx$$

which proves the statement.

#6

This problem has a simple answer: choose E_a as the event when X is *not equal* to *a*. It is not such a surprising example, as we are dealing with an *uncountable* intersection. In fact, if you take complements, the condition requires that the *union* of the complements of these sets (each having probability zero) have probability 1. This can obviously happen only because we are dealing with an uncountable collection of events.

This example reminded me of a different, more surprising, one, which you might feel has some relation to it (or maybe you won't feel that way - no matter: it is an interesting example, and it has big implications in unexpected areas, like celestial mechanics).

Consider the set of rational numbers, say, between 0 and 1. Since they are a countable set, we can write them up as a sequence, say, r_1, r_2, r_3, ... This is a *dense* set in [0,1]. Now, consider the set made up of the union of intervals of amplitude $\frac{\varepsilon}{2^n}$ around r_n, for a small ε. The total probability associated with these intervals is less than or equal to the sum of their amplitudes, which is equal to $\varepsilon \sum_n \frac{1}{2^n} = \varepsilon$. This is surprising, because ε can be as small as you wish, but the sets we are talking about are all open and dense, so that, in a topological sense, they are "big". The intersection over a sequence of such sets, as $\varepsilon \to 0$, will have probability zero. Topologically, though, it will be the intersection of a countable number of open dense sets - this is called a "second category Baire set", which, again, topologically is thought to be a "big" set (the intersection of second category Baire sets, is, again, a second category Baire set). The conclusion of this example is that "probabilistic" (or "measure theoretic") size is very different from ""'topological" size.

#10

We don't need to be this lazy, but, thanks to a quick change of variables (it is a standard one: $x \mapsto \frac{x - EX}{\sqrt{\mathrm{Var}(X)}}$), we might as well prove that $f(x) = e^{-\frac{x^2}{2}}$ has an inflection point at $|x| = 1$.
Since the second derivative of $f(x)$ is $(x^2 - 1)e^{-\frac{x^2}{2}}$ this is pretty obvious

#15

A simple way to show this is to evaluate

$$P[cX > x] = P[X > \tfrac{x}{c}] = e^{-\lambda\frac{x}{c}} = e^{-\frac{\lambda}{c}x}$$

#16

The hazard rate for an absolutely continuous random variable X is defined as the density formally defined by $\frac{P[X \in dx \mid X > x]}{dx}$, which, with a little care, can be seen to translate, in the case of absolutely continuous random variables, into $\frac{f_X(x)}{1 - F_X(x)}$, as discussed on page 234 (section 5.5.1). It is a very intuitive tool in survival and reliability theory. For a uniform variable on $[0,\, a]$ (notice that, for distributions with a continuous density, it makes no difference if we consider open, closed, semi-open, etc. intervals), this translates to

$$\frac{\frac{1}{a}}{\frac{a-x}{a}} = \frac{1}{a - x}$$

Not surprisingly, it explodes as x approaches a: after all our gadget *has to fail* at some point between 0 and a, and as we approach the end of the interval "time for failure is running out"!

#23

The density of a Weibull distribution can be defined from

$$R(t) = P[X > t] = e^{-at^b}$$

(In reliability/survival analysis we will always have $t \geq 0$). While there are approximating argument to justify this choice, the most obvious reason is seen by comparing with the exponential distribution, where

$$P[X > t] = e^{-at}$$

As the question implies, choosing values of b less than, equal to, to greater than 1, corresponds to having hazard rates that are decreasing (a "rejuvenation" effect), constant ("memoryless"), or increasing ("aging" effect). Thus, the Weibull family provides an easy toolbox for a quick and dirty description of the different age-related scenarios for the likelihood of an upcoming failure. The hazard rate calculation is an exercise in derivation:

$$\frac{f_X(t)}{R(t)} = \frac{-R'(t)}{R(t)} = -\frac{d\log R(t)}{dt} = -\frac{d\left(-at^b\right)}{dt} = abt^{b-1}$$

The conclusion about the different behavior as $b > 1$, $b = 1$, $b < 1$ follows now easily

#27

The cumulative distribution function for X is $P[X \le x] = \frac{x-a}{b-a}$. Let's define $Y = \alpha X + \beta$. Since $\alpha a + \beta \le Y \le \alpha b + \beta$, we need to make sure that $\alpha a + \beta = 0$, and $\alpha b + \beta = 1$. Thus, we want $\alpha = \frac{1}{b-a}$, $\quad \beta = \frac{a}{a-b}$. Checking, for example, the cumulative distribution function will confirm this is he right choice:

$$P[Y \le y] = P\left[X \le \frac{y-\beta}{\alpha}\right] = P[X \le a-(a-b)y] = y$$

(for $0 \le y \le 1$), by directly substituting.

#29

Note that F is a monotone increasing function. Since we are assuming that it is *strictly* monotone increasing function (and, one could be careful and extend the argument to the non strictly monotone case), it has a proper inverse F^{-1} (in the non strict case, one can use the "right inverse"). Then,

$$P[F(X) \le y] = P\left[X \le F^{-1}(y)\right] = F\left(F^{-1}(y)\right) = y$$

This observation is sometimes used to simulate a random variable with a given distribution function, starting from a program that simulates a uniform random variable: starting from a uniform random variable Y, the random variable $X = F^{-1}(Y)$ has distribution function F. This method is easy, whenever F has an easy inverse e.g., for exponential variables, but not for normal variables), but has its own numerical pitfalls - for example, simulating an exponential variable requires the computation of a logarithm, which is computer intensive, as well as less precise than algebraic calculations.

#31

We can compute the probability density function in various ways. A plain route is the following:

$$P\left[e^X \le y\right] = P[X \le \log y] = \frac{1}{\sqrt{2\pi\sigma^2}}\int_{-\infty}^{\log y} e^{-\frac{(x-\mu)^2}{2\sigma^2}}dx$$

Taking the derivative with respect to y, we obtain

$$f_Y(y) = \frac{e^{-\frac{(\log y-\mu)^2}{2\sigma^2}}}{y\sqrt{2\pi\sigma^2}}$$

This choice for stock price modeling has the advantage of simplicity, since the logarithm of the random variable has a convenient normal distribution, but it does have a number of drawbacks. One that got a lot of attention in the wake of the recent financial crisis is its lack of "fat tails" (outliers are very unlikely), but there are other concerns as well as to its appropriateness.